

OLLIVIER-RICCI CURVATURE AND THE SPECTRUM OF THE NORMALIZED GRAPH LAPLACE OPERATOR

FRANK BAUER, JÜRGEN JOST, AND SHIPING LIU

ABSTRACT. We prove the following estimate for the spectrum of the normalized Laplace operator Δ on a finite graph G ,

$$1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}}, \quad \forall \text{ integers } t \geq 1.$$

Here $k[t]$ is a lower bound for the Ollivier-Ricci curvature on the neighborhood graph $G[t]$ (here we use the convention $G[1] = G$), which was introduced by Bauer-Jost. In particular, when $t = 1$ this is Ollivier's estimate $k \leq \lambda_1$ and a new sharp upper bound $\lambda_{N-1} \leq 2 - k$ for the largest eigenvalue. Furthermore, we prove that for any G when t is sufficiently large, $1 > (1 - k[t])^{\frac{1}{t}}$ which shows that our estimates for λ_1 and λ_{N-1} are always nontrivial and the lower estimate for λ_1 improves Ollivier's estimate $k \leq \lambda_1$ for all graphs with $k \leq 0$. By definition neighborhood graphs possess many loops. To understand the Ollivier-Ricci curvature on neighborhood graphs, we generalize a sharp estimate of the curvature given by Jost-Liu to graphs which may have loops and relate it to the relative local frequency of triangles and loops.

1. INTRODUCTION

In this paper, we utilize techniques inspired by Riemannian geometry and the theory of stochastic processes in order to control eigenvalues of graphs. In particular, we shall quantify the deviation of a (connected, undirected, weighted, finite) graph G from being bipartite (a bipartite graph is one without cycles of odd lengths; equivalently, its vertex set can be split into two classes such that edges can be present only between vertices from different classes) in terms of a spectral gap. The operator whose spectrum we shall consider here is the normalized graph Laplacian Δ (it is unitarily equivalent to the one studied in Chung [8]). This is the operator underlying random walks on graphs, and so, this leads to a natural connection with the theory of stochastic processes. We observe that on a bipartite graph, a random walker, starting at a vertex x at time 0 and at each step hopping to one of the neighbors of the vertex where it currently sits, can revisit x only at even times. This connection then will be explored via the eigenvalues of Δ . More precisely, the largest eigenvalue λ_{N-1} of Δ is 2 iff G is bipartite and is < 2 else. Therefore, $2 - \lambda_{N-1}$ quantifies the deviation of G from being bipartite, and we want to understand this aspect in more detail. In more general terms, we are asking for a quantitative connection between the geometry (of the graph G) and the analysis (of the operator Δ , or the random walk encoded by it). Now, such connections have been explored systematically in Riemannian geometry, and many eigenvalue estimates are known there that connect the corresponding Laplace operator with the geometry of the underlying space M , see e.g. Li-Yau [14], Chavel [4]. The crucial role here is played by the Ricci curvature of M . In recent years, a kind of axiomatic approach to curvature has been developed. This approach encodes the abstract formal properties of curvature and thereby makes

the notion extendible to spaces more general than Riemannian manifolds. By now, there exist many notions of generalized curvature, and several of them have found important applications, see Sturm [20, 21], Lott-Villani [16], Ollivier [18], Ohta [17], Bonciocat-Sturm [2], Joulin-Ollivier [12] and the references therein. The curvature notion that turns out to be most useful for our purposes is the one introduced by Ollivier [18]. One of our main results says that

$$(1.1) \quad \lambda_{N-1} \leq 2 - k$$

where k is a lower bound for Ollivier's Ricci curvature. This matches quite well with Ollivier's result that the smallest nonzero eigenvalue λ_1 satisfies

$$(1.2) \quad \lambda_1 \geq k.$$

In fact, one of the main points of the present paper is to relate such upper and lower bounds via random walks. In fact, as in Bauer-Jost [1], we translate this relationship into the geometric concept of a neighborhood graph. The idea here is that in the t -th neighborhood graph $G[t]$ of G , vertices x and y are connected by an edge with a weight given by the probability that a random walker starting at x reaches y after t steps times the degree of x . We note that even though the original graph may have been unweighted, the neighborhood graphs $G[t]$ are necessarily weighted. In addition, they will in general possess self-loops, because the random walker starting at x may return to x after t steps. Therefore, we need to develop our theory on weighted graphs with self-loops even though the original G might have been unweighted and without such loops. Since Ollivier's curvature is defined in terms of transportation distances (Wasserstein metrics), we can then use our neighborhood graphs in order to geometrically control the transportation costs and thereby to estimate the curvature of the neighborhood graphs in terms of the curvature of the original graph. As it turns out that lower bounds for the smallest eigenvalue of $G[2]$ are related to upper bounds for the largest eigenvalue of G , we can then use methods for controlling the smallest eigenvalue to also obtain (1.1). For controlling the smallest eigenvalue, besides Ollivier [18], we also refer to Lin-Yau [15] and Jost-Liu [11]. In particular, in the last paper, we could relate λ_1 to the local clustering coefficient introduced in Watts-Strogatz [24]. The local clustering coefficient measures the relative local frequency of triangles, that is, cycles of length 3. Since bipartite graphs cannot possess any triangles, this then is obviously related to our question about quantifying the deviation of the given graph G from being bipartite. In fact, in Jost-Liu [11], this local clustering has been controlled in terms of Ollivier's Ricci curvature. Thus, in the present paper we are closing the loop between the geometric properties of a graph G , the spectrum of its graph Laplacian, random walks on G , and the generalized curvature of G , drawing upon deep ideas and concepts originally developed in Riemannian geometry and the theory of stochastic processes.

2. THE NORMALIZED LAPLACE OPERATOR, NEIGHBORHOOD GRAPHS, AND OLLIVIER-RICCI CURVATURE

In this paper, $G = (V, E)$ will denote an undirected, weighted, connected, finite graph of N vertices. We do not exclude loops, i.e., we permit the existence of an edge between a vertex and itself. V denotes the set of vertices and E denotes the set of edges. If two vertices $x, y \in V$ are connected by an edge, we say x and y are neighbors, in symbols $x \sim y$. The associated weight function $w: V \times V \rightarrow \mathbb{R}$ satisfies $w_{xy} = w_{yx}$ (because the graph is undirected) and we assume $w_{xy} > 0$ whenever $x \sim y$.

and $w_{xy} = 0$ iff $x \not\sim y$; thus, we do not permit negative weights. For a vertex $x \in V$, its degree d_x is defined as $d_x := \sum_{y \in V} w_{xy}$. If $w_{xy} = 1$ whenever $x \sim y$, we shall call the graph an unweighted one. We will also speak of a locally finite graph $\tilde{G} = (\tilde{V}, \tilde{E})$, which is an undirected, weighted, connected graph and satisfies the property that for every $x \in \tilde{V}$, the number of edges connected to x is finite.

2.1. The normalized graph Laplace operator and its eigenvalues. In this subsection, we recall the definition of the normalized graph Laplace operator and state some of its basic properties. In particular, we will emphasize the relations between eigenvalues of the Laplace operator and random walks on graphs.

Let $C(V)$ denote the space of all real-valued functions on the set V . We attach to each vertex $x \in V$ a probability measure $m_x(\cdot)$ and then can define the Laplace operator.

Definition 1. *The Laplace operator Δ is defined as*

$$(2.1) \quad \Delta f(x) = \sum_{y \in V} f(y)m_x(y) - f(x), \quad \forall f \in C(V).$$

The measure $m_x(\cdot)$ can also be considered as the distribution of a 1-step random walk starting from x . We will choose

$$(2.2) \quad m_x(y) = \begin{cases} \frac{w_{xy}}{d_x}, & \text{if } y \sim x, \\ 0, & \text{otherwise,} \end{cases}$$

in the following. Note that $x \sim x$ is possible when x has a loop. With this family of probability measures $\{m_x(\cdot)\}$, Δ is just the normalized graph Laplace operator studied in Grigoryan [10] and Bauer-Jost [1] and is unitarily equivalent to the Laplace operator studied in Chung [8].

Remark 1. *On a graph G without loops, we can also consider a lazy random walk. A lazy random walk is a random walk that does not move with a certain probability, i.e. for some x we might have $m_x(x) \neq 0$. In this case, the lazy random walk on G is equivalent to the usual random walk on the graph G^{lazy} that is obtained from G by adding for every vertex x a loop with a weight $d_x m_x(x)$.*

We also have a natural measure μ on the whole set V ,

$$\mu(x) := d_x,$$

which gives an inner product structure on $C(V)$.

Definition 2. *The inner product of two functions $f, g \in C(V)$ is defined as*

$$(2.3) \quad (f, g)_\mu = \sum_{x \in V} f(x)g(x)\mu(x).$$

Then $C(V)$ becomes a Hilbert space, and we can write $C(V) = l^2(V, \mu)$. By the definition of the degree and the symmetry of the weight function, we can check that

- μ is invariant w.r.t. $\{m_x(\cdot)\}$, i.e. $\sum_{x \in V} m_x(y)\mu(x) = \mu(y)$, $\forall y \in V$;
- μ is reversible w.r.t. $\{m_x(\cdot)\}$, i.e. $m_x(y)\mu(x) = m_y(x)\mu(y)$, $\forall x, y \in V$.

These two facts imply immediately that the operator Δ is nonpositive and self-adjoint on the space $l^2(V, \mu)$. We call λ an eigenvalue of Δ if there exists some $f \neq 0$ such that

$$(2.4) \quad \Delta f = -\lambda f.$$

Using this convention it follows from the observation that Δ is self-adjoint and non-positive that all its eigenvalues are real and nonnegative. In fact, it's well known that (see e.g. Chung [8])

$$(2.5) \quad 0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 2.$$

Since our graph is connected we actually have $0 < \lambda_1$. In Chung [8] it is shown, by proving a discrete version of the Cheeger inequality, that λ_1 is a measure for how easy/difficult it is to cut the graph into two large pieces. Furthermore, $\lambda_{N-1} = 2$ if and only if G is bipartite. In Bauer-Jost [1] it is shown, by establishing a dual version of the Cheeger inequality for the largest eigenvalue, that λ_{N-1} is a measure of the size of a locally bipartite subgraph that is connected by only a few edges to the rest in the graph. In the following, we will call λ_1 the first eigenvalue and λ_{N-1} the largest eigenvalue of the operator Δ .

2.2. Neighborhood graphs. In this subsection, we recall the neighborhood graph method developed by Bauer-Jost [1].

As discussed above, the Laplace operator underlies random walks on graphs. In this section, we discuss the deep relationship between eigenvalues estimates for the Laplace operator Δ and random walks on the graph G by using neighborhood graphs.

We first introduce the following notation. For a probability measure μ , we denote

$$\mu P(\cdot) := \sum_x \mu(x) m_x(\cdot).$$

Let δ_x be the Dirac measure at x , then we can write $\delta_x P^1(\cdot) := \delta_x P(\cdot) = m_x(\cdot)$. Therefore the distribution of a t -step random walk starting from x with a transition probability m_x is

$$(2.6) \quad \delta_x P^t(\cdot) = \sum_{x_1, \dots, x_{t-1}} m_x(x_1) m_{x_1}(x_2) \cdots m_{x_{t-1}}(\cdot)$$

for $t > 1$. In particular, using the measure (2.2), the probability that the random walk starting at x moves to y in t steps is given by

$$\delta_x P^t(y) = \begin{cases} \sum_{x_1, \dots, x_{t-1}} \frac{w_{xx_1}}{d_x} \frac{w_{x_1 x_2}}{d_{x_1}} \cdots \frac{w_{x_{t-1} y}}{d_{x_{t-1}}}, & \text{if } t > 1; \\ \frac{w_{xy}}{d_x}, & \text{if } t = 1. \end{cases}$$

The idea is now to define a family of graphs $G[t]$, $t \geq 1$ that encodes the transition probabilities of the t -step random walks on the graph G .

Definition 3. *The neighborhood graph $G[t] = (V, E[t])$ of the graph $G = (V, E)$ of order $t \geq 1$ has the same vertex set as G and the weights of the edges of $G[t]$ are defined by the transitions probabilities of the t -step random walk,*

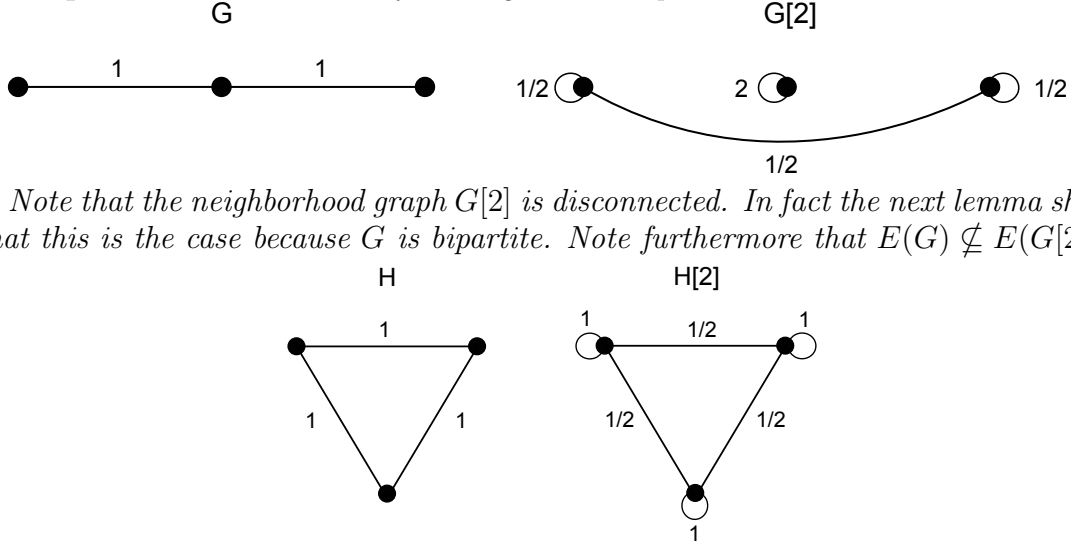
$$(2.7) \quad w_{xy}[t] := \delta_x P^t(y) d_x.$$

In particular, $G = G[1]$ and $x \sim y$ in $G[t]$ if and only if there exists a path of length t between x and y in G .

Remark 2. *We note here that the neighborhood graph method is related to the discrete heat kernel $p_t(x, y)$ on graphs. For more details about the discrete heat kernel see for instance Grigoryan [10]. We have*

$$p_t(x, y) = \frac{w_{xy}[t]}{d_x d_y}.$$

Example 1. We consider the following two examples.



Note that the neighborhood graph $G[2]$ is disconnected. In fact the next lemma shows that this is the case because G is bipartite. Note furthermore that $E(G) \not\subseteq E(G[2])$.

For this example we have $E(H) \subseteq E(H[2])$.

These examples show that the neighborhood graph $G[t]$ is in general a weighted graph with loops, even if the original graph G is an unweighted, simple graph.

Lemma 1. The neighborhood graph $G[t]$ has the following properties (see Bauer-Jost [1]):

- (i) If t is even, then $G[t]$ is connected if and only if G is not bipartite. Furthermore, if t is even, $G[t]$ cannot be bipartite.
- (ii) If t is odd, then $G[t]$ is always connected and $G[t]$ is bipartite iff G is bipartite.
- (iii) $d_x[t] = d_x$ for all $x \in V$.

Note that (iii) implies that $l^2(V, \mu) = l^2(V, \mu[t])$ for all t . In particular, we have the same inner product for all neighborhood graphs $G[t]$. The crucial observation is the next theorem.

Theorem 1 (Bauer-Jost). The Laplace operator Δ on G and the Laplace operator $\Delta[t]$ on $G[t]$ are related to each other by the following identity:

$$\text{id} - (\text{id} - \Delta)^t = \Delta[t].$$

The nonlinear relation between Δ and $\Delta[t]$ leads directly to the following eigenvalue estimates.

Theorem 2 (Bauer-Jost). Let $\mathcal{A}[t]$ be a lower bound for the eigenvalue $\lambda_1[t]$ of $\Delta[t]$, i.e., $\lambda_1[t] \geq \mathcal{A}[t]$. Then

$$(2.8) \quad 1 - (1 - \mathcal{A}[t])^{\frac{1}{t}} \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 1 + (1 - \mathcal{A}[t])^{\frac{1}{t}}$$

if t is even and

$$(2.9) \quad 1 - (1 - \mathcal{A}[t])^{\frac{1}{t}} \leq \lambda_1$$

if t is odd.

Theorem 3 (Bauer-Jost). Let $\mathcal{B}[t]$ be an upper bound for the largest eigenvalue $\lambda_{N-1}[t]$ of $\Delta[t]$, i.e. $\lambda_{N-1}[t] \leq \mathcal{B}[t]$. (Since $\lambda_{N-1}[t] \leq 1$ for t even, we can assume in this case w.l.o.g. that $\mathcal{B}[t] \leq 1$ in this case.) Then all eigenvalues of Δ are contained in the union of the intervals

$$\left[0, 1 - (1 - \mathcal{B}[t])^{\frac{1}{t}}\right] \cup \left[1 + (1 - \mathcal{B}[t])^{\frac{1}{t}}, 2\right]$$

if t is even and

$$\lambda_{N-1} \leq 1 - (1 - \mathcal{B}[t])^{\frac{1}{t}}$$

if t is odd.

These theorems show how random walks on graphs (or equivalently neighborhood graphs) can be used to estimate eigenvalues of the Laplace operator. In the rest of this paper we will use these insights to derive lower bounds for λ_1 and upper bounds for λ_{N-1} in terms of the Ollivier-Ricci curvature of a graph.

2.3. Ollivier-Ricci curvature from a probabilistic view. We define a metric d on the set of vertices V as follows. For two distinct points $x, y \in V$, $d(x, y)$ is the number of edges in the shortest path connecting x and y . Then, including the family of probability measures $m := \{m_x(\cdot)\}$, we have a structure (V, d, m) , on which the definition of Ricci curvature proposed by Ollivier [18] can be stated.

Definition 4 (Ollivier). *For any two distinct points $x, y \in V$, the (Ollivier-) Ricci curvature of (V, d, m) along (xy) is defined as*

$$(2.10) \quad \kappa(x, y) := 1 - \frac{W_1(m_x, m_y)}{d(x, y)}.$$

Here, $W_1(m_x, m_y)$ is the transportation distance between the two probability measures m_x and m_y , in a formula,

$$(2.11) \quad W_1(m_x, m_y) = \inf_{\xi^{x,y} \in \Pi(m_x, m_y)} \sum_{x' \in V} \sum_{y' \in V} d(x', y') \xi^{x,y}(x', y'),$$

where $\Pi(m_x, m_y)$ is the set of probability measures $\xi^{x,y}$ that satisfy

$$(2.12) \quad \sum_{y' \in V} \xi^{x,y}(x', y') = m_x(x'), \quad \sum_{x' \in V} \xi^{x,y}(x', y') = m_y(y').$$

The conditions (2.12) simply ensure that we start with the measure m_x and end up with m_y . Intuitively, $W_1(m_x, m_y)$ is the minimal cost to transport the mass of m_x to that of m_y with the distance as the cost function. We also call such a $\xi^{x,y}$ a transfer plan between m_x and m_y , or a coupling of two random walks governed by m_x and m_y respectively. Those $\xi^{x,y}$ ($\xi^{x,y}$ might not be unique) which attain the infimum value in (2.11), are called optimal couplings. The optimal coupling exists in a very general setting. For locally finite graphs the existence follows from a simple and interesting argument in Remark 14.2 in [13].

There is a Kantorovich duality formula for transportation distances,

$$(2.13) \quad W_1(m_x, m_y) = \sup_{f: \text{Lip}(f) \leq 1} \left[\sum_{x' \in V} f(x') m_x(x') - \sum_{y' \in V} f(y') m_y(y') \right],$$

where $\text{Lip}(f) := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}$. For more details about this concept, we refer to Villani [22, 23], and Evans [9].

Remark 3. *Applying this definition of curvature to a Riemannian manifold with a family of probability measures obtained by restricting the volume measure to the closed r -balls, κ will reduce to the Ricci curvature as $r \rightarrow 0$ up to a scaling r^2 , see Ollivier [18], for which we also refer for a general treatment of Ricci curvature. In Riemannian geometry, Ricci curvature controls how fast the geodesics emanating from the same point diverge on average, or equivalently how fast the volume of distance balls grows*

as a function of the radius. Jost-Liu [11] translate those ideas into a combinatorial setting and show that Ollivier-Ricci curvature on a locally finite graph reflects the relative abundance of triangles, which is captured by the local clustering coefficient introduced by Watts-Strogatz [24].

For the rest of this paper, let k be a lower bound for the Ollivier-Ricci curvature, i.e.

$$(2.14) \quad \kappa(x, y) \geq k, \quad \forall x \sim y.$$

Remark 4. By Proposition 19 in Ollivier [18], it follows that if k is a lower curvature bound for all neighbors x, y then it is a lower curvature bound for all pairs of vertices. This also follows from Theorem 4 below.

Remark 5. By definition, the lower bound k for the curvature κ is no larger than one. In fact, such a lower bound k always exists. Since the largest possible distance between points from the supports of two measures m_x and m_y at a pair of neighbors x, y is 3, we can easily estimate $\kappa(x, y) \geq -2$. We will derive a more precise lower bound for κ on a locally finite graph with loops in Theorem 6, see also Lin-Yau [15] and Jost-Liu [11] for related results.

We could also write (2.14) as

$$(2.15) \quad W_1(m_x, m_y) \leq (1 - k)d(x, y) = 1 - k, \quad \forall x \sim y,$$

which is essentially equivalent to the well known path coupling criterion on the state space of Markov chains used to study the mixing time of them (see Bubley-Dyer [3] or [13], [19]). We will utilize this idea to interpret the lower bound of the Ollivier-Ricci curvature as a control on the expectation value of the distance between two coupled random walks.

We reformulate Bubley-Dyer's theorem (see [3] or [13], [19]) in our language.

Theorem 4 (Bubley-Dyer). *On (V, d, m) , if for each pair of neighbors $x, y \in V$, we have the contraction*

$$W_1(m_x, m_y) \leq (1 - k)d(x, y) = 1 - k,$$

then for any two probability measures μ and ν on V , we have

$$W_1(\mu P, \nu P) \leq (1 - k)W_1(\mu, \nu).$$

With this at hand, it is easy to see that if for any pair of neighbors x, y , $\kappa(x, y) \geq k$, then for any time t and any two \bar{x}, \bar{y} , which are not necessarily neighbors, the following is true,

$$(2.16) \quad W_1(\delta_{\bar{x}}P^t, \delta_{\bar{y}}P^t) \leq (1 - k)^t d(\bar{x}, \bar{y}).$$

We consider two coupled discrete time random walks (\bar{X}_t, \bar{Y}_t) , whose distributions are $\delta_{\bar{x}}P^t, \delta_{\bar{y}}P^t$ respectively. They are coupled in a way that the probability

$$p(\bar{X}_t = \bar{x}', \bar{Y}_t = \bar{y}') = \xi_t^{\bar{x}, \bar{y}}(\bar{x}', \bar{y}'),$$

where $\xi_t^{\bar{x}, \bar{y}}(\cdot, \cdot)$ is the optimal coupling of $\delta_{\bar{x}}P^t$ and $\delta_{\bar{y}}P^t$. In this language, we can interpret the term $W_1(\delta_{\bar{x}}P^t, \delta_{\bar{y}}P^t)$ as the expectation value of the distance $\mathbf{E}^{\bar{x}, \bar{y}}d(\bar{X}_t, \bar{Y}_t)$ between the coupled random walks \bar{X}_t and \bar{Y}_t .

Corollary 1. *On (V, d, m) , if $\kappa(x, y) \geq k, \forall x \sim y$, then we have for any two $\bar{x}, \bar{y} \in V$,*

$$(2.17) \quad \mathbf{E}^{\bar{x}, \bar{y}}d(\bar{X}_t, \bar{Y}_t) = W_1(\delta_{\bar{x}}P^t, \delta_{\bar{y}}P^t) \leq (1 - k)^t d(\bar{x}, \bar{y}).$$

2.4. Ollivier-Ricci curvature and the lower bound of the first eigenvalue.

In his paper [18], Ollivier proved a spectral gap estimate which works on a general metric space with random walks. In particular, on finite graphs, it could be stated as follows.

Theorem 5 (Ollivier). *On (V, d, m) , if $\kappa(x, y) \geq k$, $\forall x \sim y$, then the first eigenvalue λ_1 of the normalized graph Laplace operator Δ satisfies*

$$\lambda_1 \geq k.$$

This is a discrete analogue of the estimate for the smallest nonzero eigenvalue of the Laplace-Beltrami operator on a Riemannian manifold by Lichnerowicz. As pointed out in Ollivier [18], this result is also related to the coupling method for estimates of the first eigenvalue in the Riemannian setting developed by Chen-Wang [7] (which leads to a refinement of the eigenvalue estimate of Li-Yau [14]), see also the surveys Chen [5, 6]. The corresponding result of Corollary 1 in the smooth case, i.e., controlling the expectation distance of two coupled Markov chains in terms of the lower bound of Ricci curvature on a Riemannian manifold, is a key step in Chen-Wang's method.

A direct proof of Theorem 5 can be found in [18]. Here we want to present an analogue of Chen-Wang's method in the discrete setting. Our proof of Theorem 5 motivated us to combine the Ollivier-Ricci curvature and the neighborhood graph method via random walks. It reflects the deep connection between eigenvalue estimates and random walks or heat equations.

Proof: We consider the transition probability operator $P : l^2(V, \mu) \rightarrow l^2(V, \mu)$ defined by $Pf(x) := \sum_y f(y)m_x(y) = \sum_y f(y)\delta_x P(y)$. Then we have $P^t f(x) = \sum_y f(y)\delta_x P^t(y)$. We construct a discrete time heat equation,

$$(2.18) \quad \begin{cases} f(x, 0) = f_1(x), \\ f(x, 1) - f(x, 0) = \Delta f(x, 0), \\ f(x, 2) - f(x, 1) = \Delta f(x, 1), \\ \dots \\ f(x, t+1) - f(x, t) = \Delta f(x, t), \end{cases}$$

where $f_1(x)$ satisfies $\Delta f_1(x) = -\lambda_1 f_1(x) = Pf_1(x) - f_1(x)$. Iteratively, one can find the solution of the above system of equations as

$$(2.19) \quad f(x, t) = P^t f_1(x) = (1 - \lambda_1)^t f_1(x).$$

We remark here that the solution of the heat equation on a Riemannian manifold with the first eigenfunction as the initial value is $f(x, t) = f_1(x)e^{-\lambda_1 t}$, which also involves information about both the eigenvalue λ_1 and the eigenfunction $f_1(x)$.

If $\lambda_1 \geq 1$ there is nothing to prove since $\kappa \leq 1$ by definition. Therefore we can suppose $\lambda_1 < 1$ in the following. Then we have for any $\bar{x}, \bar{y} \in V$

$$\begin{aligned} (1 - \lambda_1)^t |f_1(\bar{x}) - f_1(\bar{y})| &= |f(\bar{x}, t) - f(\bar{y}, t)| \\ &= |P^t f_1(\bar{x}) - P^t f_1(\bar{y})| \\ &\leq \sum_{\bar{x}', \bar{y}'} |f(\bar{x}') - f(\bar{y}')| \xi_t^{\bar{x}, \bar{y}}(\bar{x}', \bar{y}') \\ &\leq \text{Lip}(f_1) \mathbf{E}^{\bar{x}, \bar{y}} d(\bar{X}_t, \bar{Y}_t) \\ &\leq \text{Lip}(f_1) (1 - k)^t d(\bar{x}, \bar{y}). \end{aligned}$$

Here, $\text{Lip}(f)$ is always finite since the underlying space V is a finite set. In the last inequality we used Corollary 1. From an analytic point of view, the above calculation can be seen as a gradient estimate for the solution of the heat equation.

Since the eigenfunction f_1 for the eigenvalue λ_1 is orthonogal to the constant function, i.e. $(f_1, \mathbf{1})_\mu = 0$, we can always find $x_0, y_0 \in V$ such that $|f_1(x_0) - f_1(y_0)| > 0$. It follows that

$$0 < \left(\frac{1-k}{1-\lambda_1} \right)^t \text{Lip}(f_1) d(x_0, y_0).$$

To prevent a contradiction when $t \rightarrow \infty$, we need

$$(2.20) \quad \frac{1-k}{1-\lambda_1} \geq 1,$$

which completes the proof. \square

3. ESTIMATES FOR OLLIVIER-RICCI CURVATURE ON LOCALLY FINITE GRAPHS WITH LOOPS

Jost-Liu [11] give a sharp estimate for Ollivier-Ricci curvature on locally finite graphs without loops. As shown in Example 1, neighborhood graphs are in general weighted graphs with loops. Therefore, for our purposes, we need to understand the curvature of graphs with loops. In this section, we generalize the estimates in Jost-Liu [11] for locally finite graphs $\tilde{G} = (\tilde{V}, \tilde{E})$ that may have loops. This is done by considering a novel optimal transportation plan.

We first fix some notations. For any two real numbers a, b ,

$$a_+ := \max\{a, 0\}, \quad a \wedge b := \min\{a, b\}, \quad \text{and} \quad a \vee b := \max\{a, b\}.$$

We denote $\tilde{N}_x := \{z \in \tilde{V} | z \sim x\}$ as the neighborhood of x and $N_x := \tilde{N}_x \cup \{x\}$. Then $N_x = \tilde{N}_x$ if x has a loop. For every pair of neighbors x, y , we divide N_x, N_y into disjoint parts as follows.

$$(3.1) \quad N_x = \{x\} \cup \{y\} \cup N_x^1 \cup N_{xy}, \quad N_y = \{y\} \cup \{x\} \cup N_y^1 \cup N_{xy},$$

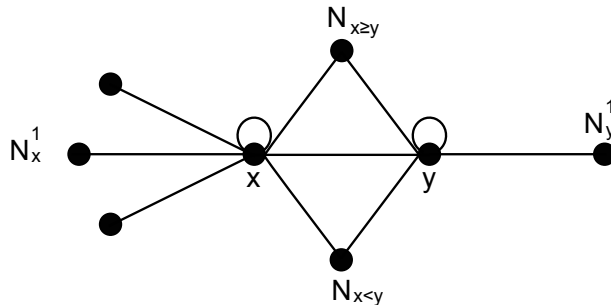
where

$$N_{xy} = N_{x \geq y} \cup N_{x < y}$$

and

$$\begin{aligned} N_x^1 &:= \{z | z \sim x, z \not\sim y, z \neq y\}, \\ N_{x \geq y} &:= \{z | z \sim x, z \sim y, z \neq x, z \neq y, \frac{w_{xz}}{d_x} \geq \frac{w_{zy}}{d_y}\}, \\ N_{x < y} &:= \{z | z \sim x, z \sim y, z \neq x, z \neq y, \frac{w_{xz}}{d_x} < \frac{w_{zy}}{d_y}\}. \end{aligned}$$

In the next figure we illustrate this partition of the vertex set.



Theorem 6. On $\tilde{G} = (\tilde{V}, \tilde{E})$, we have for any pair of neighbors $x, y \in \tilde{V}$,

$$\begin{aligned} \kappa(x, y) \geq k(x, y) := & - \left(1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1x}}{d_x} \vee \frac{w_{x_1y}}{d_y} \right)_+ \\ & - \left(1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1x}}{d_x} \wedge \frac{w_{x_1y}}{d_y} \right)_+ \\ & + \sum_{x_1 \in N_{xy}} \frac{w_{x_1x}}{d_x} \wedge \frac{w_{x_1y}}{d_y} + \frac{w_{xx}}{d_x} + \frac{w_{yy}}{d_y}. \end{aligned}$$

Moreover, this inequality is sharp.

Remark 6. On an unweighted graph, the form $k(x, y)$ for $x \sim y$ becomes

$$k(x, y) = - \left(1 - \frac{1}{d_x} - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_x \wedge d_y} \right)_+ - \left(1 - \frac{1}{d_x} - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_x \vee d_y} \right)_+ + \frac{\sharp(x, y)}{d_x \vee d_y} + \frac{c(x)}{d_x} + \frac{c(y)}{d_y},$$

where $\sharp(x, y) := \sum_{x_1 \in N_{xy}} 1$ is the number of triangles containing x, y , $c(x) = 0$ or 1 is the number of loops at x .

Proof: Since the total mass of m_x is equal to one, we obtain from (3.1) the following identity for neighboring vertices x and y :

$$(3.2) \quad 1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1x}}{d_x} = \frac{w_{xx}}{d_x} + \sum_{x_1 \in N_x^1} \frac{w_{x_1x}}{d_x}$$

A similar identity holds for y .

We denote

$$\begin{aligned} A_{x,y} &:= 1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1x}}{d_x} \vee \frac{w_{x_1y}}{d_y}, \\ B_{x,y} &:= 1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1x}}{d_x} \wedge \frac{w_{x_1y}}{d_y}. \end{aligned}$$

Obviously, $A_{x,y} \leq B_{x,y}$. We firstly try to understand these two quantities.

If $A_{x,y} \geq 0$, we have

$$(3.3) \quad 1 - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1y}}{d_y} \geq \frac{w_{xy}}{d_x} + \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1y}}{d_y} \right),$$

i.e., using (3.2) we observe that the mass of m_y at y and N_y^1 is no smaller than that of m_x at y and the excess mass at $N_{x \geq y}$. Rewriting (3.3) in the form

$$\frac{w_{xy}}{d_y} + \sum_{x_1 \in N_{xy}} \frac{w_{x_1y}}{d_y} \leq 1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1y}}{d_y} \right),$$

and subtracting the term $\sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y}$ on both sides we obtain

$$(3.4) \quad \frac{w_{xy}}{d_y} + \sum_{x_1 \in N_{x < y}} \left(\frac{w_{x_1y}}{d_y} - \frac{w_{xx_1}}{d_x} \right) \leq 1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x},$$

i.e., the mass of m_x at x and N_x^1 is larger than that of m_y at x and the excess mass at $N_{x < y}$.

If $B_{x,y} \geq 0$, we have

$$(3.5) \quad 1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} + \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1y}}{d_y} \right) \geq \frac{w_{xy}}{d_y},$$

i.e., the mass of m_x at x and N_x^1 and the excess mass at $N_{x \geq y}$ is no smaller than that of m_y at x .

In Jost-Liu [11] it is explicitly described how much mass has to be moved from a vertex in N_x to which point in N_y , i.e. the exact value of $\xi^{x,y}(x', y')$, for any $x' \in N_x$, $y' \in N_y$. But in the case with loops it would be too complicated if we try to do the same thing. Instead, we adopt here a dynamic strategy. That is, we think of a discrete time flow of mass. After one unit time, the mass flows forward for distance 1 or stays there. We only need to determine the direction of the flow according to different cases.

As in Jost-Liu [11], we divide the discussion into 3 cases.

- $0 \leq A_{x,y} \leq B_{x,y}$. In this case we use the following transport plan: Suppose the initial time is $t = 0$.

$t = 1$: Move all the mass at N_x^1 to x and the excess mass at $N_{x \geq y}$ to y . We denote the distribution of the mass after the first time step by m^1 . We have

$$W_1(m_x, m^1) \leq \left(1 - \frac{w_{xx}}{d_x} - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \right) \times 1 + \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1y}}{d_y} \right) \times 1$$

$t = 2$: Move one part of the excess mass at x now to fill the gap at $N_{x < y}$ and the other part to y . By (3.4) the mass at x after $t = 1$ is enough to do so. The distribution of the mass is now denoted by m^2 . We have

$$\begin{aligned} W_1(m^1, m^2) &\leq \sum_{x_1 \in N_{x < y}} \left(\frac{w_{x_1y}}{d_y} - \frac{w_{xx_1}}{d_x} \right) \times 1 \\ &\quad + \left[\left(1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \right) - \sum_{x_1 \in N_{x < y}} \left(\frac{w_{x_1y}}{d_y} - \frac{w_{xx_1}}{d_x} \right) - \frac{w_{xy}}{d_y} \right] \times 1 \end{aligned}$$

$t = 3$: Move the excess mass at y now to N_y^1 . We denote the mass after the third time step by $m^3 = m_y$. We have

$$\begin{aligned} W_1(m^2, m_y) &\leq \left[\left(1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \right) - \sum_{x_1 \in N_{x < y}} \left(\frac{w_{x_1y}}{d_y} - \frac{w_{xx_1}}{d_x} \right) - \frac{w_{xy}}{d_y} + \frac{w_{xy}}{d_x} \right. \\ &\quad \left. + \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1y}}{d_y} \right) - \frac{w_{yy}}{d_y} \right] \times 1 \end{aligned}$$

By triangle inequality and (2.11), we get

$$\begin{aligned} W_1(m_x, m_y) &\leq W_1(m_x, m^1) + W_1(m^1, m^2) + W_1(m^2, m_y) \\ &= 3 - 2\frac{w_{xy}}{d_x} - 2\frac{w_{xy}}{d_y} - 2 \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \vee \frac{w_{x_1y}}{d_y} - \frac{w_{xx}}{d_x} - \frac{w_{yy}}{d_y}. \end{aligned}$$

Moreover, if the following function can be extended as a function on the graph such that $\text{Lip}(f) \leq 1$, (i.e., if there are no paths of length 1 between N_x^1 and

$N_{x<y}$, nor paths of length 1 between N_y^1 and $N_{x\geq y}$, nor paths of length 1 or 2 between N_x^1 and N_y^1),

$$f(z) = \begin{cases} 0, & \text{if } z \in N_y^1; \\ 1, & \text{if } z \in \{y\} \cup N_{x<y}; \\ 2, & \text{if } z \in \{x\} \cup N_{x\geq y}; \\ 3, & \text{if } z \in N_x^1, \end{cases}$$

then by Kantorovich duality (2.13), we can show that the inequality above is actually an equality. Recalling the definition of $\kappa(x, y)$, we have proved the theorem in this case.

- $A_{x,y} < 0 \leq B_{x,y}$. We use the following transfer plan:

$t = 1$: We divide the excess mass of m_x at $N_{x\geq y}$ into two parts. One part together with the mass of m_x at y is enough to fill gaps at y and N_y^1 . Since (3.3) doesn't hold in this case, this is possible. We move this part of mass to y and the other part to x . We also move all the mass of m_x at N_x^1 to x .

$t = 2$: We move the excess mass at x now to $N_{x<y}$ and the excess mass at y to N_y^1 .

Applying this transfer plan, we can prove (we omit the calculation here)

$$W_1(m_x, m_y) \leq 2 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - 2 \sum_{x_1 \in N_{xy}} \left(\frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y} \right) - \frac{w_{xx}}{d_x} - \frac{w_{yy}}{d_y}.$$

Moreover, if the following function can be extended as a function on the graph such that $\text{Lip}(f) \leq 1$, (i.e., if there are no paths of length 1 between $N_x^1 \cup N_{x\geq y}$ and $N_y^1 \cup N_{x<y}$),

$$f(z) = \begin{cases} 0, & \text{if } z \in N_y^1 \cup N_{x<y}; \\ 1, & \text{if } z = x \text{ or } z = y; \\ 2, & \text{if } z \in N_x^1 \cup N_{x\geq y}, \end{cases}$$

then by Kantorovich duality (2.13), we can check that the inequality above is actually an equality.

- $A_{x,y} \leq B_{x,y} < 0$. We use the following transport plan:

$t = 1$: Move the mass of m_x at N_x^1 and $N_{x\geq y}$ to x . Since now (3.5) doesn't hold, we need to move one part of the mass $m_y(y)$ to x and the other part to N_y^1 and $N_{x<y}$.

Applying this transfer plan, we can calculate

$$W_1(m_x, m_y) \leq 1 - \sum_{x_1 \in N_{xy}} \left(\frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y} \right) - \frac{w_{xx}}{d_x} - \frac{w_{yy}}{d_y}.$$

Since the following function can be extended as a function on the graph such that $\text{Lip}(f) \leq 1$,

$$f(z) = \begin{cases} 0, & \text{if } z \in \{x\} \cup N_{x<y} \cup N_y^1; \\ 1, & \text{if } z \in \{y\} \cup N_{x\geq y} \cup N_x^1, \end{cases}$$

we can check the inequality above is in fact an equality by Kantorovich duality. That is, in this case for any $x \sim y$,

$$\kappa(x, y) = \sum_{x_1 \in N_{xy}} \left(\frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y} \right) + \frac{w_{xx}}{d_x} + \frac{w_{yy}}{d_y}.$$

□

We also have a generalization of the upper bound in Jost-Liu [11] on \tilde{G} .

Theorem 7. *On $\tilde{G} = (\tilde{V}, \tilde{E})$, we have for every pair of neighbors x, y ,*

$$\kappa(x, y) \leq \sum_{x_1 \in \{x\} \cup \{y\} \cup N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y}.$$

Proof: $I := \sum_{x_1 \in \{x\} \cup \{y\} \cup N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y}$ is exactly the mass of m_x which we need not move. The other mass need to be moved for at least distance 1. So we have $W_1(m_x, m_y) \geq 1 - I$, which implies $\kappa(x, y) \leq I$, for $x \sim y$. □

Remark 7. *If for every $x \in \tilde{V}$, $w_{xx} = 0$, then Theorem 6 and Theorem 7 reduce to the estimates in Jost-Liu [11].*

Remark 8. *Theorem 7 tells us that if $\kappa(x, y) > 0$ for any $x \sim y$, then either $\sharp(x, y) \neq 0$ or $c(x) \neq 0$ or $c(y) \neq 0$.*

Example 2. *We consider a lazy random walk on an unweighted complete graph \mathcal{K}_N with N vertices governed by $m_x(y) = 1/N, \forall x, y$. Or equivalently, we consider the graph $\mathcal{K}_N^{\text{lazy}}$. Using Theorem 6 and Theorem 7, we get for any x, y*

$$1 = \frac{N-2}{N} + \frac{1}{N} + \frac{1}{N} \leq \kappa(x, y) \leq \frac{1}{N} \cdot N = 1.$$

That is, in this case, both the lower and the upper bound are sharp.

4. ESTIMATES OF THE SPECTRUM IN TERMS OF OLLIVIER-RICCI CURVATURE

In this section, we use the neighborhood graph method to derive estimates for the spectrum of the normalized graph Laplace operator in terms of the Ollivier-Ricci curvature. In particular, we get a sharp upper bound for the largest eigenvalue using the curvature on the original graph. Furthermore, we give nontrivial estimates using the curvature on neighborhood graphs even if the curvature of the original graph is nonpositive. In this case, our results improve Theorem 5.

Lemma 2. *Let $k[t]$ be a lower bound of Ollivier-Ricci curvature of the neighborhood graph $G[t]$. Then for all $t \geq 1$ the eigenvalues of Δ on G satisfy*

$$(4.1) \quad 1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}}$$

if t is even and

$$(4.2) \quad 1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1$$

if t is odd. Moreover, if G is not bipartite, then there exists a $t' \geq 1$ such that for all $t \geq t'$ the eigenvalues of Δ on G satisfy

$$0 < 1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}} < 2$$

if t is even and

$$0 < 1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1$$

if t is odd.

Proof: By combining Theorem 2 and Theorem 5 we directly obtain equation (4.1) and equation (4.2).

The second part of this Lemma is proved in two steps. In the first step, we will show that if G is not bipartite then there exists a t' such that for all $t \geq t'$ the neighborhood

graph $G[t]$ of G satisfies $w_{xy}[t] \neq 0$ for all $x, y \in V$, i.e. $G[t]$ is a complete graph and each vertex has a loop. In the second step, we show that any graph that satisfies $w_{xy} \neq 0$ for all $x, y \in V$ has a positive lower curvature bound, i.e. $k > 0$. This then completes the proof.

Step 1: By the definition of the neighborhood graph it is sufficient to show that for all $t \geq t'$ there exists a path of length t between any pair of vertices. Since G is not bipartite it follows that there exists a path of even and a path of odd length between any pair of vertices in the graph (This is the definition of a bipartite graph!). Given a path of length L between x and y then we can find a path of length $L + 2$ between x and y as follows: We go in L steps from x to y and then from y to one of its neighbors and then back to y . This is a path of length $L + 2$ between x and y . Since G is finite, it follows that there exists a t' such that there are paths of length t for all $t \geq t'$ for any pair of vertices.

Step 2: Given a graph that satisfies $w_{xy} \neq 0$ for all $x, y \in V$.

Since each vertex in the graph is a neighbor of all other vertices, it is clear that we can move the excess mass of m_x for distance 1 to anywhere. Therefore

$$W_1(m_x, m_y) \leq 1 - \sum_{x_1 \in V} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y},$$

which implies

$$\kappa(x, y) \geq \sum_{x_1 \in V} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y}.$$

By Theorem 7, it follows that the above inequality is in fact an equality. Hence for all $x, y \in V$, we have

$$\kappa(x, y) = \sum_{x_1 \in V} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y} \geq N \frac{\min_{x,y} w_{xy}}{\max_x d_x} \geq \frac{\min_{x,y} w_{xy}}{\max_{x,y} w_{xy}} > 0,$$

since the weight is positive for every pair of vertices.

This completes the proof. \square

Remark 9. From Remark 8 after Theorem 7 we know that positive lower curvature bound is a strong restriction on a graph. Hence, Olliver's estimate $\lambda_1 \geq k$ only yields nontrivial estimates for λ_1 in those restricted cases. Lemma 2 shows that, unless G is bipartite, the neighborhood graph method can be used to always obtain a nontrivial lower bound for λ_1 in terms of the Ollivier-Ricci curvature.

Now, we will show that it is also possible to control the largest eigenvalue of Δ on the graph G in terms of the Ollivier-Ricci curvature of G itself.

Theorem 8. On (V, d, m) , if $\kappa(x, y) \geq k$, $\forall x \sim y$, then the largest eigenvalue λ_{N-1} of the normalized graph Laplace operator Δ satisfies

$$\lambda_{N-1} \leq 2 - k.$$

Remark 10. As discussed above, the largest eigenvalue λ_{N-1} is small if the graph does not contain a locally bipartite subgraph that has only a few connections to the rest of the graph. It is well known that a graph is bipartite if and only if it does not contain any cycle of odd length, in particular a bipartite graph does not contain triangles and loops. Since positive Ollivier-Ricci curvature implies that locally there always exists a triangle or a loop, i.e. the graph cannot be locally bipartite, it is not surprising that the largest eigenvalue can actually be controlled from above in terms of the Ollivier-Ricci curvature.

By combining Theorem 8 and Theorem 6, we have the following corollary. From Remark 6, we know that $k(x, y)$ reflects the number of triangles and loops in the graph.

Corollary 2. *On $G = (V, E)$, the largest eigenvalue satisfies*

$$\lambda_{N-1} \leq 2 - \min_{x \sim y} k(x, y),$$

where $k(x, y)$ is defined in Theorem 6.

The main point in the proof of Theorem 8 is to explore the relation of the Ollivier-Ricci curvature κ on the original graph G and the Ollivier-Ricci curvature $\kappa[t]$ on its neighborhood graph $G[t]$. Before we prove Theorem 8 we consider some lemmata.

If we interpret the graph $G = (V, E)$ as a structure $(V, d, m = \{\delta_x P\})$, then by (2.7) its neighborhood graph $G[t] = (V, E[t])$ can be considered as a structure $(V, d[t], \{\delta_x P^t\})$. So the first step should be to estimate the variance of the metrics on neighborhood graphs.

Lemma 3. *For any $x, y \in V$, we have*

$$(4.3) \quad \frac{1}{t} d(x, y) \leq d[t](x, y).$$

Proof: For any $x, y \in V$, we set $d[t](x, y) = \infty$ if we cannot find a path connecting them in $G[t]$. Otherwise, we just choose a shortest path $x_0 = x, x_1, \dots, x_l = y$, between x and y in $G[t]$, i.e. $l = d[t](x, y)$. For x_i, x_{i+1} , $i = 0, \dots, l-1$, by definition of neighborhood graph, we have $d(x_i, x_{i+1}) \leq t$ in G . Equivalently,

$$\frac{1}{t} d(x_i, x_{i+1}) \leq 1 = d[t](x_i, x_{i+1}).$$

Summing over all i , we get

$$\frac{1}{t} \sum_{i=0}^{l-1} d(x_i, x_{i+1}) \leq d[t](x, y).$$

Then the triangle inequality of d on G gives (4.3). □

Remark 11. *In fact, when t is larger than the diameter D of the graph G , we have a better estimate*

$$(4.4) \quad \frac{1}{t} d(x, y) \leq \frac{1}{D} d(x, y) \leq 1 \leq d[t](x, y).$$

Lemma 4. *If $E \subseteq E[t]$, then $d[t](x, y) \leq d(x, y)$.*

The proof is obvious. The interesting point is that when the Ollivier-Ricci curvature of the graph G is positive, $E \subseteq E[t]$ is satisfied for all t and hence Lemma 4 is applicable. This can be seen as follows. First we observe that for $(x, y) \in E$, if $\sharp(x, y) \neq 0$ or $c(x) \neq 0$ or $c(y) \neq 0$, then $(x, y) \in E[t]$. Then we get the conclusion immediately from Remark 8 after Theorem 7.

Lemma 5. *Let k be a lower bound of κ on G . If $E \subseteq E[t]$, then the curvature $\kappa[t]$ of the neighborhood graph $G[t]$ satisfies*

$$(4.5) \quad \kappa[t](x, y) \geq 1 - t(1 - k)^t, \quad \forall x, y \in V.$$

Proof: By Lemma 4, Corollary 1 and Lemma 3, we get

$$\begin{aligned} W_1^{d[t]}(\delta_x P^t, \delta_y P^t) &\leq W_1^d(\delta_x P^t, \delta_y P^t) \\ &\leq (1-k)^t d(x, y) \\ &\leq t(1-k)^t d[t](x, y). \end{aligned}$$

We use $W_1^{d[t]}$, W_1^d here to indicate the different cost functions used in these two quantities. In the first inequality above we used that the transportation distance (2.11) is linear in the graph distance $d(\cdot, \cdot)$. Recalling the definition of the curvature, we have proved (4.5). \square

Now we arrive at the point to prove the upper bound of the largest eigenvalue. **Proof of Theorem 8:** Using Lemma 5 and Theorem 5, we know on $G[t]$,

$$\lambda_1[t] \geq 1 - t(1-k)^t.$$

Then by Lemma 2, we get for any even number t ,

$$\lambda_{N-1} \leq 1 + t^{\frac{1}{t}}(1-k).$$

Letting $t \rightarrow +\infty$, we get $\lambda_{N-1} \leq 2 - k$. \square

By combining Theorem 5 and Theorem 8, we get

$$k \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 2 - k,$$

That is, (4.1) is also true for $t = 1$.

Example 3. On an unweighted complete graph \mathcal{K}_N with N vertices, we have

$$k = \kappa(x, y) = \frac{N-2}{N-1}, \forall x, y \text{ and } \lambda_1 = \dots = \lambda_{N-1} = \frac{N}{N-1}.$$

Therefore,

$$k < \lambda_1 = \dots = \lambda_{N-1} = 2 - k.$$

That is, our upper bound estimate for λ_{N-1} is sharp for unweighted complete graphs.

Example 4. Let's revisit the graph $\mathcal{K}_N^{\text{lazy}}$ in Example 2. We have

$$k = \kappa(x, y) = 1, \forall x, y \text{ and } \lambda = \dots = \lambda_{N-1} = 1.$$

Therefore,

$$k = \lambda_1 = \dots = \lambda_{N-1} = 2 - k = 1.$$

That is, both estimates are sharp in this case.

In fact, we can show that (4.1) is true for any $t \geq 1$.

Theorem 9. Let $k[t]$ be a lower bound of Ollivier-Ricci curvature of the neighborhood graph $G[t]$. Then for all $t \geq 1$ the eigenvalues of Δ on G satisfy

$$(4.6) \quad 1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}}.$$

Moreover, if G is not bipartite, then there exists a $t' \geq 1$ such that for all $t \geq t'$ the eigenvalues of Δ on G satisfy

$$0 < 1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}} < 2.$$

Proof: By Lemma 2, we only need to prove the upper estimate $\lambda_{N-1} \leq 1 + (1 + k[t])^{\frac{1}{t}}$ if t is odd. This follows immediately from Theorem 8 and Theorem 3. \square

In particular, unless G is bipartite, we can obtain nontrivial lower bounds for λ_1 and nontrivial upper bounds for λ_{N-1} even if $k = k[1]$ is nonpositive.

5. ESTIMATES FOR THE LARGEST EIGENVALUE IN TERMS OF THE NUMBER OF JOINT NEIGHBORS

In Bauer-Jost [1] it is shown that the next lemma is a simple consequence of Theorem 1.

Lemma 6. *Let u be an eigenfunction of Δ for the eigenvalue λ . Then,*

$$(5.1) \quad 2 - \lambda = \frac{(u, \Delta[2]u)_\mu}{(u, \Delta u)_\mu} = \frac{\sum_{x,y} w_{xy}[2](u(x) - u(y))^2}{\sum_{x,y} w_{xy}(u(x) - u(y))^2}.$$

Lemma 6 can be used to derive further estimates for the largest eigenvalue λ_{N-1} from above and below. We introduce the following notations:

Definition 5. *Let \tilde{N}_x be the neighborhood of vertex x as in Section 3. The minimal and the maximal number of joint neighbors of any two neighboring vertices is defined as $\tilde{\sharp}_1 := \min_{x \sim y} (\sharp(x, y) + c(x) + c(y))$ and $\tilde{\sharp}_2 := \max_{x \sim y} (\sharp(x, y) + c(x) + c(y))$, respectively. Furthermore, we define $W := \max_{x,y} w_{xy}$ and $w := \min_{x,y; x \sim y} w_{xy}$.*

Theorem 10. *We have the following estimates for λ_{N-1} :*

(i) *If $E(G) \subseteq E(G[2])$ then*

$$\lambda_{N-1} \leq 2 - \frac{w^2}{W} \frac{\tilde{\sharp}_1}{\max_x d_x}.$$

(ii) *If $E(G[2]) \subseteq E(G)$ then*

$$2 - \frac{W^2}{w} \frac{\tilde{\sharp}_2}{\min_x d_x} \leq \lambda_{N-1}$$

Proof. On the one hand, we observe that if $E(G) \subseteq E(G[2])$ then for every pair of neighboring vertices $x \sim y$ in G

$$(5.2) \quad \frac{w_{xy}[2]}{w_{xy}} = \frac{\sum_z \frac{1}{d_z} w_{xz} w_{zy}}{w_{xy}} \geq \frac{w^2}{W} \frac{\tilde{\sharp}_1}{\max_x d_x}.$$

On the other hand if $E(G[2]) \subseteq E(G)$ then for every pair of neighboring vertices $x \sim y$ in $G(2)$ we have

$$(5.3) \quad \frac{w_{xy}[2]}{w_{xy}} = \frac{\sum_z \frac{1}{d_z} w_{xz} w_{zy}}{w_{xy}} \leq \frac{W^2}{w} \frac{\tilde{\sharp}_2}{\min_x d_x}.$$

Substituting the inequalities (5.2) and (5.3) in equation (5.1) completes the proof. \square

Interestingly, Theorem 10 (i) yields an alternative proof of Theorem 8 in the case of unweighted regular graphs. Since Theorem 8 trivially holds if $k \leq 0$ we only have to consider the case when $k > 0$ is a lower curvature bound. The discussion after Lemma 4 shows that $k > 0$ implies that $E(G) \subseteq E(G[2])$ and hence we can apply Theorem 10 (i) in this case. From Theorem 7 it follows that for an unweighted graph

$$\kappa(x, y) \leq \frac{\sharp(x, y)}{d_x \vee d_y} + \frac{c(x)}{d_x} + \frac{c(y)}{d_y}$$

for all pairs of neighboring vertices x, y . In the case of an d -regular graph G this implies that a lower bound k for the Ollivier-Ricci curvature must satisfy,

$$k \leq \frac{\tilde{\sharp}_1}{d}.$$

Hence for an unweighted d -regular graph Theorem 10 implies

$$\lambda_{N-1} \leq 2 - \frac{\tilde{\sharp}_1}{d} \leq 2 - k.$$

We consider the following example.

Example 5. *Let's revisit the complete unweighted graph \mathcal{K}_N . For the complete graph we have $E(\mathcal{K}_N) \subseteq E(\mathcal{K}_N[2])$. We have $\tilde{\sharp}_1 = N - 2$ and $\max_x d_x = N - 1$. Thus, Theorem 10 (i) yields*

$$\lambda_{N-1} \leq \frac{N}{N-1},$$

i. e. the estimate from above is sharp for complete graphs. Now we consider the unweighted complete graph on N vertices with N loops $\mathcal{K}_N^{\text{lazy}}$. We have $E(\mathcal{K}_N^{\text{lazy}}[2]) = E(\mathcal{K}_N^{\text{lazy}})$. Furthermore, we have $\tilde{\sharp}_2 = N$ and $\min_x d_x = N$. Thus, Theorem 10 (ii) yields

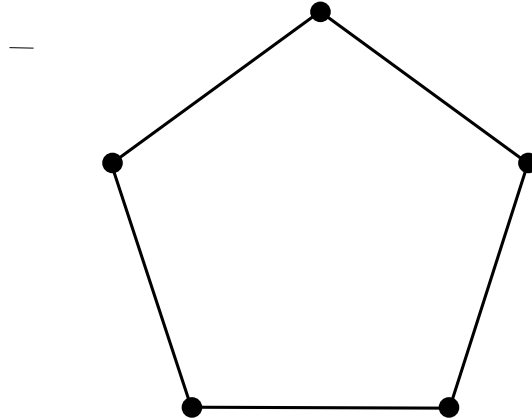
$$1 \leq \lambda_{N-1},$$

i. e. the estimate from below is sharp for complete graphs with loops.

6. AN EXAMPLE

In this section, we explore a particular example, the circle \mathcal{C}_5 with 5 vertices. We show that our estimates using the neighborhood graph method can yield nontrivial estimates although the curvature of the original graph has a nonpositive lower curvature bound. We also discuss the growth rate of the lower bound k for curvature $\kappa[t]$ as $t \rightarrow \infty$ on \mathcal{C}_5 .

We consider the unweighted graph \mathcal{C}_5 displayed in the next figure.

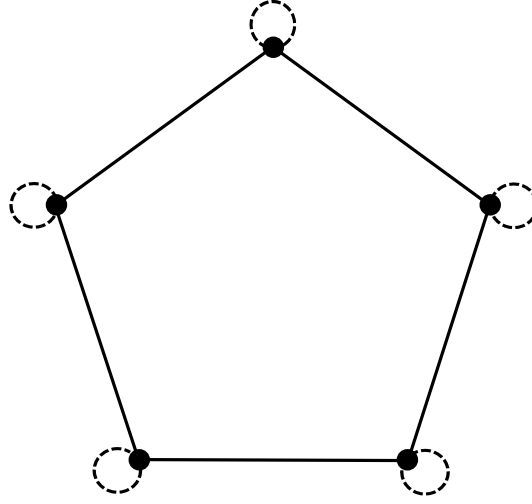


We know that the first and largest eigenvalue of Δ on \mathcal{C}_5 are given by

$$\lambda_1 = 1 - \cos \frac{2\pi}{5} \doteq 0.6910, \quad \lambda_4 = 1 - \cos \frac{4\pi}{5} \doteq 1.8090.$$

It is easy to check that the optimal lower bound k for the curvature is 0. So in this case Ollivier's result Theorem 5 and Theorem 8 only yield trivial estimates.

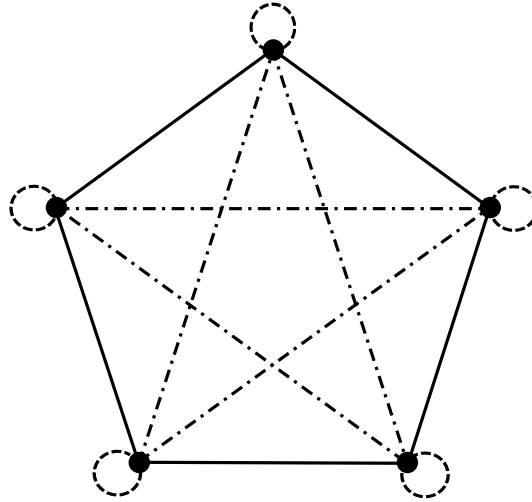
Now we consider the neighborhood graph $\mathcal{C}_5[2]$ depicted in the next figure (we change the order of vertices).



The weight of every dashed loop is 1 and the weight of every solid edge is $1/2$. We can check that the optimal $k[2]$ is $1/4$. Then Theorem 9 yields the nontrivial estimate,

$$\lambda_1 \geq 1 - \frac{\sqrt{3}}{2} \doteq 0.1340, \quad \lambda_4 \leq 1 + \frac{\sqrt{3}}{2} \doteq 1.8660.$$

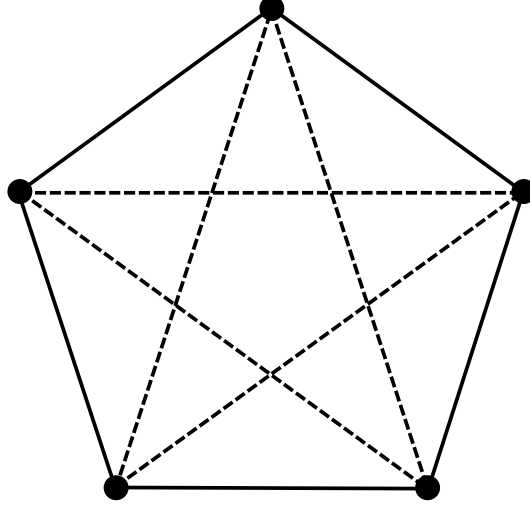
Moreover, the neighborhood graph $\mathcal{C}_5[4] = (\mathcal{C}[2])[2]$ is depicted in the next figure.



The weight of every dashed loop is $3/4$ and the weight of every solid edge is $1/2$ and every dash-dotted edge is $1/8$. We can check that the optimal lower curvature bound is $k[4] = 1/2$. So Theorem 9 tells us that

$$\lambda_1 \geq 1 - \frac{1}{\sqrt[4]{2}} \doteq 0.1591, \quad \lambda_4 \leq 1 + \frac{1}{\sqrt[4]{2}} \doteq 1.8409.$$

We can also consider the neighborhood graph of odd order $\mathcal{C}_5[3]$, which is depicted in the next figure.



The weight of every solid edge is $3/4$ and the weight of every dashed edge is $1/4$. Then the optimal lower curvature bound is given by $k[3] = 3/8$. Theorem 9 implies

$$\lambda_1 \geq 1 - \left(\frac{5}{8}\right)^{\frac{1}{3}} \doteq 0.1450, \quad \lambda_4 \leq 1 + \left(\frac{5}{8}\right)^{\frac{1}{3}} \doteq 1.8550.$$

From the above calculations, we see that the Ollivier-Ricci curvature $\kappa[t]$ on the neighborhood graphs give better and better estimates in this example.

We also further compare the estimates in Theorem 10 and Theorem 8 on the weighted graph $\mathcal{C}_5[3]$. By Theorem 10 we get

$$\lambda_4[3] \leq 2 - \frac{(1/4)^2}{3/4} \cdot \frac{3}{2} = \frac{15}{8}.$$

This is worse than the one given by Theorem 8, $\lambda_4[3] \leq 2 - \frac{3}{8} = \frac{13}{8}$.

Another interesting problem is concerned with the limit of the neighborhood graphs. As shown in Bauer-Jost [1], since \mathcal{C}_5 is a regular non-bipartite graph, $\mathcal{C}_5[t]$ will converge to $\mathcal{C}_5[\infty] := \bar{\mathcal{C}}_5$ as $t \rightarrow \infty$. In this case $\bar{\mathcal{C}}_5$ is a complete graph and every vertex has a loop and the weight of every edge in $\bar{\mathcal{C}}_5$ is $2/5$. We can then check that $\kappa[\infty] = 1$. By Theorem 9, for large enough t , we have

$$(6.1) \quad 0 < 1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 = 1 - \cos \frac{2\pi}{5} < 1.$$

Therefore $0 < (1 - k[t])^{\frac{1}{t}} < 1$. Intuitively, $k[t]$ should become larger and larger as $t \rightarrow \infty$ since the graph $\mathcal{C}_5[t]$ converges to a complete graph and its weights become more and more the same. We suppose $\lim_{t \rightarrow \infty} (1 - k[t])^{\frac{1}{t}}$ exists (this is true at least for a subsequence). Then to avoid contradictions in (6.1), we know there exists a positive number a such that

$$\lim_{t \rightarrow \infty} (1 - k[t])^{\frac{1}{t}} = e^{-a} > 0.$$

That is,

$$\lim_{t \rightarrow \infty} \frac{\log(1 - k[t])}{t} = -a,$$

which means $k[t]$ behaves like $1 - P(t)e^{-at}$ as $t \rightarrow \infty$ where $P(t)$ is a polynomial in t .

REFERENCES

- [1] F. Bauer and J. Jost, *Bipartite and neighborhood graphs and the spectrum of the normalized graph Laplacian*, <http://arxiv.org/abs/0910.3118v3>, to appear in Comm. Anal. Geom..
- [2] A.-I. Bonciocat and K.-T. Sturm, *Mass transportation and rough curvature bounds for discrete spaces*, J. Funct. Anal. **256** (2009), no. 9, 2944-2966.
- [3] R. Bubley and M. E. Dyer, *Path coupling: A technique for proving rapid mixing in Markov chains*, 38th Annual Symposium on Foundations of Computer Science (FOCS '97), (1997), 223-231.
- [4] I. Chavel, *Eigenvalues in Riemannian geometry*, Academic Press, 1984.
- [5] M. Chen, *Trilogy of couplings and general formulas for lower bound of spectral gap*, Probability towards 2000 (New York, 1995), 123-136, Lecture Notes in Statist., **128**, Springer, New York, 1998.
- [6] M. Chen, *Coupling, spectral gap and related topics I*, Chinese Sci. Bull. **42** (1997), no. 16, 1321-1327.
- [7] M. Chen and F. Wang, *Application of coupling method to the first eigenvalue on manifold*, Sci. China Ser. A **37** (1994), no. 1, 1-14.
- [8] F. R. K. Chung, *Spectral graph theory*, CBMS Regional Conference Series in Mathematics, **92**, 1997.
- [9] L. C. Evans, *Partial differential equations and Monge-Kantorovich mass transfer*, Current developments in mathematics, 1997 (Cambridge, MA), 65-126, Int. Press, Boston, MA, 1999.
- [10] A. Grigoryan, *Analysis on graphs*, Lecture notes University of Bielefeld, 2009.
- [11] J. Jost and S. Liu, *Ollivier's Ricci curvature, local clustering and curvature dimension inequalities on graphs*, <http://arxiv.org/pdf/1103.4037v2>.
- [12] A. Joulin and Y. Ollivier, *Curvature, concentration and error estimates for Markov chain Monte Carlo*, Ann. Probab. **38** (2010), no. 6, 2418-2442.
- [13] D. A. Levin, Y. Peres and E. L. Wilmer, *Markov chains and mixing times*. With a chapter by James G. Propp and David B. Wilson. American Mathematical Society, Providence, RI, 2009.
- [14] P. Li and S. T. Yau, *Estimates of eigenvalues of a compact Riemannian manifold*, AMS Proc. Symp. Pure Math. **36** (1980), 205-240.
- [15] Y. Lin and S. T. Yau, *Ricci curvature and eigenvalue estimate on locally finite graphs*, Math. Res. Lett. **17** (2010), no. 2, 343-356.
- [16] J. Lott and C. Villani, *Ricci curvature for metric measure spaces via optimal transport*, Ann. of Math. **169** (2009), no. 3, 903-991.
- [17] S.-I. Ohta, *On the measure contraction property of metric measure spaces*, Comment. Math. Helv. **82** (2007), no. 4, 805-828.
- [18] Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. **256** (2009), no. 3, 810-864.
- [19] Y. Peres, *Mixing for Markov chains and spin systems*, Lecture notes, <http://www.stat.berkeley.edu/~peres/ubc.pdf>, 2005.
- [20] K.-T. Sturm, *On the geometry of metric measure spaces. I.*, Acta Math. **196** (2006), no. 1, 65-131.
- [21] K.-T. Sturm, *On the geometry of metric measure spaces. II.*, Acta Math. **196** (2006), no. 1, 133-177.
- [22] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, **58**. American Mathematical Society, Providence, RI, 2003.
- [23] C. Villani, *Optimal transport, Old and new*, Grundlehren der Mathematischen Wissenschaften, **338**. Springer-Verlag, Berlin, 2009.
- [24] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks*. Nature **393** (1998), 440-442.

MAX PLANCK INSTITUTE FOR MATHEMATICS IN THE SCIENCES, LEIPZIG 04103, GERMANY.
E-mail address: `Frank.Bauer@mis.mpg.de`

MAX PLANCK INSTITUTE FOR MATHEMATICS IN THE SCIENCES, LEIPZIG 04103, GERMANY.
E-mail address: `jost@mis.mpg.de`

MAX PLANCK INSTITUTE FOR MATHEMATICS IN THE SCIENCES, LEIPZIG 04103, GERMANY
AND ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES,
BEIJING 100190, CHINA.
E-mail address: `shiping@mis.mpg.de`